Face Recognition Using Contourlet Transform under Varying Illumination

Elham Hashemi Shad[1,2], Sedigheh Ghofrani[1,*]

[1]Electronic and Electrical Engineering Department, Islamic Azad University, South Tehran Branch, Tehran, Iran
[2]Department of Management Systems, Quality and Inspection, Standard Research Institute, Karaj, Iran

Elham_hashemi_shad@yahoo.com; s_ghofrani@azad.ac.ir
*Corresponding Author

Abstract

This paper has presented an algorithm for face recognition under varying illumination by using Contourlet transform (CT) as a two dimensional transform defined in discrete form and principal component analysis (PCA) as a linear subspace. The Contourlet coefficients of images were obtained for multiple scales and directions then projected to PCA subspace to obtain feature vectors. This feature vectors were projected once again to the PCA and the output was used for classification. We used twice PCA to reduce overloading of information and extraction of the key features. All of the Contourlet frequency coefficients were required to overcome the illumination drastic changes. In our experiments, to improve the processing precision, we considered 3 levels of decomposition in three scales and a total of 15 orientations along with the low pass subband image. Our method relatively is robust against of varying illumination. In this work, the Euclidean distance as the simplest classifier was used for classification. To evaluate our algorithm, we tested it on the Extended Yale B face database and showed the appropriate performance in comparison with other approaches.

Keywords

Face Recognition; Contourlet Transform; Principal Component Analysis; Singular Value Decomposition; Euclidean Distance

Introduction

Human face recognition is an important issue in many applications such as public security, human computer interface (Myers, 1998), and human authentication (Carstens et al., 2004). These applications range from static matching of controlled format photograph (e.g. passport, credit card and driving license) to real time matching of intelligent video images surveillance (Stonham, 1986). The main challenges for any face recognition system are pose, illumination, age and facial expression variation and also face exclusion by hair, scarf and glasses. Face recognition under varying illumination is a challenging issue because of changing the face appearance dramatically with varying illumination. In fact, changes due to illumination may be greater than those due to face identity. In this paper, variations were taken into consideration due to illumination alone. Many researches have been performed to overcome this problem; of which Kim et al. (2008) presented a preprocessing method using the Census transform that generates a number of binary values from the current pixel by comparion with the neighbor pixels, and encoding the binary values into a single gray scale image and then applied the principal component analysis (PCA) to the gray scale images. It is worth nothing that the Census transform is a non-parametric transform first introduced by Zabih et al. (1996). This transform compared the intensity of the center pixel with neighbour pixels and then it produced a series of binary numbers. In fact, these binary numbers are local features; Hai-Long et al. (2010) proposed a Contourlet multi-threshold method for uneven illumination face images. The proposed algorithm combines hard threshold with two dimensional (2-D) shadow compensation method and the selection of proper thresholds depends on the subband layers of Contourlet transform. Logarithmic nonsubsampled Contourlet transform (LNSCT) to estimate the reflectance component from a single face image referring to the illumination invariant feature for face recognition was proposed in (Xie et al., 2010). One of the techniques for face recognition is correlation, where the test image is recognized by the closest point in the training set so that all of the images have zero mean and unit variance. In fact, all images are normalized (Georghiades et al., 2001). In general, linear subspace method, an approach for face recognition for model of the varying illumination, has the 3-D linear subspace that the distance of the test image is computed to each linear subspace and the shortest distance is chosen as equivalent image (Georghiades et al., 2001). The cones-attached is a type of illumination cone for face recognition system (Georghiades et al., 2001).

In this paper, we used Contourlet transform (CT) for face recognition. For this propose, the original image was decomposed into 3 scales and 15 orientations. Then all subbands were utilized along with the low pass subband, in fact, all frequencies were used for obtaining the feature vectors. The feature vectors were obtained by using the principle component analysis (PCA) and then projected once again to the linear subspace of PCA. The Euclidean distance was used for classification. This method was relatively robust against drastic changes in lighting and in comparison with other methods has appropriate performance.

This paper was organized as follows. In Section 2, CT and PCA were expressed briefly. Our proposed method for feature extraction and face recognition was explained in Section 3. Section 4, showed the achieved experimental results of our proposed method in comparison with other approaches where the Extended Yale B face database was used. Conclusion was given in section 5.

Background

In this section, first of all, we had a short review on CT as a new multiresolution transform in image processing, then we briefly explained the PCA.

Discrete Contourlet Transform

The CT is a new image decomposition scheme which provides the flexible multi-resolution representation for two dimension signals (Minh et al., 2005). The two properties that make CT superior to other transforms such as wavelet are directionality and anisotropy (Minh et al., 2005). The block diagram of CT for two level decomposition is shown in Fig. 1 which includes two sub-blocks, laplacian pyramid (LP) (Burt et al., 1983) for multi-resolution and multi-scale representation and directional filter bank (DFB) (Bamberger et al., 1992) for multi-direction decomposition. The block diagram of LP is shown in Fig. 2 where H and G are analysis and synthesis filters, respectively and MS denotes the sampling matrix. As seen, the output of LP at each level includes the sampled low pass and the band pass version of the input signal.

The band pass image was then processed by the DFB and the same steps was repeated upon the low pass image until the multi-scale and multi-direction decomposition of an image was obtained. The DFB was implemented via an n-level tree-structured decomposition led to $d = 2^n$ multi-bands with wedge-shaped frequency partition as shown in Fig. 3.

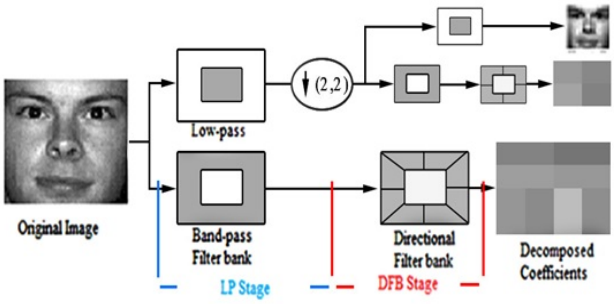

FIG. 1 THE BLOCK DIAGRAM OF TWO LEVELS CT.

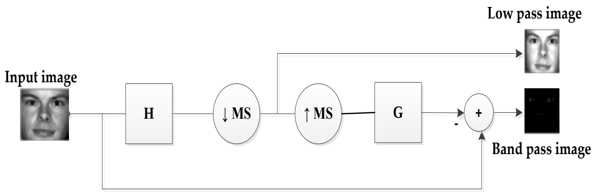

FIG. 2 THE BLOCK DIAGRAM OF LP.

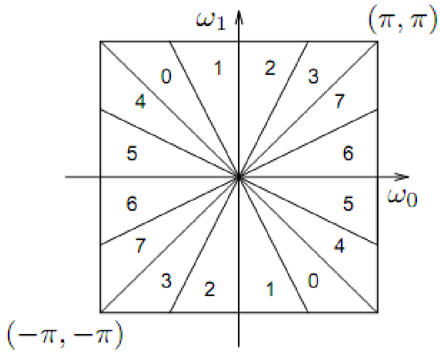

FIG. 3 THE k REAL WEDGE-SHAPE FREQUENCY BANDS OF DFB FOR n=3, AND $d = 2^3$ (MOON ET AL.,2001).

FIG. 4 CT COEFFICIENTS OF A SAMPLE FACE IMAGE FOR 3 LEVELS OF DECOMPOSITION.

Directional information was obtained by the two operator (Yang et al., 2008); a two-channel quincunx filter bank with fan filter which divides a 2-D spectrum into two directions of vertical and horizontal, and a shearing operator reordering the image pixels. In general, LP decomposes the input image into the low pass and band pass image, then, DFB decomposes the band pass image in order to capture the directional information. The same above procedures were

repeated on the low-pass image at the second level of CT. Combination of LP and DFB gave a double filter bank structure known as Contourlet filter bank. The Contourlet filter bank decomposed the given image into directional multi-bands at multiple scales. Fig. 4 shows samples of a sub-band image and Contourlet coefficients at 3 scales and 15 orientations.

Principal Component Analysis

Principal component analysis (PCA) is a well-known statistical method, often applied as the pre-processing stage of face recognition (Moon et al., 2001). In general, PCA is, first, to reduce the dimension and therefore computational complexity, and second, extract the input data important features (Mandal et al., 2007). Suppose the size of the input data for PCA is $M \times N$ where $M < N$. The procedures of PCA for the input data are written as follows,

- arrange the input data

$$A = \{A_i\} \ ; \ i = 1, 2, ..., N \tag{1}$$

where A_i denotes the original vectorized image, with size $M \times 1$ then clearly A is a matrix with size $M \times N$.

- compute the mean vector of the matrix,

$$\mu = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{2}$$

- if the datasets aren't centered, it can be centered by subtracting the mean vector, μ , from every vectors,

$$\phi_i = A_i - \mu \ , \ i = 1, 2, ..., N \tag{3}$$

Obviously, the size of each ϕ_i is $M \times 1$.

- compute the covariance matrix ,

$$C = \frac{1}{N}\sum_{i=1}^{N} \phi_i \phi_i^T \tag{4}$$

with size $M \times M$ where $(.)^T$ denotes the transpose operation.

- determine the eigen values, $\{\lambda_i\}_{i=1:M}$, and eigen vectors, $\{V_i\}_{i=1:M}$, of covariance matrix, whre $\lambda_1 > \lambda_2 > \ldots > \lambda_M$, $|C - \lambda I| = 0$, $CV_i = \lambda_i V_i$.

Eigen values and eigen vectors can be also obtained via the singular value decomposition.

- use the eigen vectors corresponding to the k largest eigen values, named feature vectors,

$$u_k = \sum_{i=1}^{k} V_i \tag{5}$$

where $k << M$ and so $k << N$.

It was shown that considering a small set of Eigen vectors corresponds to the largest Eigen values still the image characteristics are stored(Pirra et al., 2011).

- project the feature vectors to linear subspace,

$$E = U_k^T B \tag{6}$$

Where $B = \{\phi_i\}_{i=1:N}$ with size $M \times N$ and U matrix is instituted from u vectors. The above linear transform is used to reduce the dimension and classification is prepared as well.

Face Recognition based on CT

In this work, we decomposed the input image by CT into 3 levels. Therefore, the contourlet coefficients of all training images were calculated at 3 scales and 15 orientations. In fact, a total of 16 subband included 3 scales and 15 orientation along with the low pass subband image. Each subband was put into a column vector with size $P_j \times 1$, where j is the subband index. On the other hand, N images were considered as train. So, as Fig. 5 shows, the size of each matrix including N training images (under different illumination) is $P_j \times N$ where $j = 1, \ldots, 16$. At first, we considered first subbands of N training images. Therefore, we had a matrix with size of $P_1 \times N$. Then, the PCA were applied on this matrix to reduce the dimension and we obtained eigen values and eigen vectors by using singular value decomposition.

Compute the singular value decomposition to determine the Eigen vectors of C^{sd} ,

$$U^{sd} S^{sd} ((V)^{sd})^T = C^{sd} \tag{7}$$

Where s and d represent scale and orientation, respectively. C^{sd} denotes the covariance matrix of Contourlet coefficients, U^{sd} consists of Eigen vectors sorted according to the decreasing order of Eigen values and the diagonal matrix S^{sd} contains the Eigen values. In fact, the columns of U are the left singular vectors (gene coefficient vectors); S (the same dimensions as matrix of Contourlet coefficients) has singular values and is diagonal (mode amplitudes); and V^T has rows that are the right singular vectors (expression level vectors). The SVD represents an expansion of the original data in a coordinate system where the covariance matrix is diagonal. Let $\{\lambda_i\}_{i=1:N}$ represent the eigen values in decreasing order then K

is observed for the remaining 95% of the energy, the following

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \approx 0.95 \tag{8}$$

$$U_k^{sd} = U_i^{sd} \qquad \text{where} \qquad i = 1 \ldots k \tag{9}$$

In general, each time we get 95% energy of subbands (It is also significant results with 100% energy). The subspace contourlet coefficients are given by

$$B^{sd} = (U_k^{sd})^T (\phi^{sd}) \tag{10}$$

Where ϕ^{sd} denotes matrix of the datasets centered. Note that B^{sd} is obtained for a subspace at the same scale s and same orientation d, similarly, B^{sd} is obtained in the same way for all subbands.

The subspace Contourlet coefficients were normalized so that the variance along each of the L dimensions became equal. This was done by dividing the subspace coefficients by the square root of the respective eigen values. The normalized subspace Contourlet coefficients at three levels of each image were stacked to form a matrix of feature vectors B where each column was a feature vector of the concatenated subspace Contourlet coefficients of an image.

$$B = [B^{11}; B^{12}; \ldots; B^{21}; B^{22}; \ldots] \tag{11}$$

In fact, each column of B corresponded to an image and row corresponded to a specific subspace Contourlet coefficients. The above mentioned procedures are shown in Fig. 5.

These features were once again projected to a linear subspace, however, this time without subtracting the mean.

Since the feature dimension is usually larger compared to the size of the training data, BB^T is very large. Therefore, the pseudo covariance matrix $C = B^T B$ was calculated instead and the singular value decomposition was computed to determine the eigen vectors.

At the end, suppose q and t represent the query/test and train image respectively. in h dimensions, the Euclidean distance between two point t and q is,

$$d_{Euc}(t,q) = \sqrt{\sum_{i=1}^{h} (t_i - q_i)^2} \tag{12}$$

That is used for classification. t_i or q_i is the coordinate of t or q in dimension i. The above mentioned procedures are shown in Fig. 5.

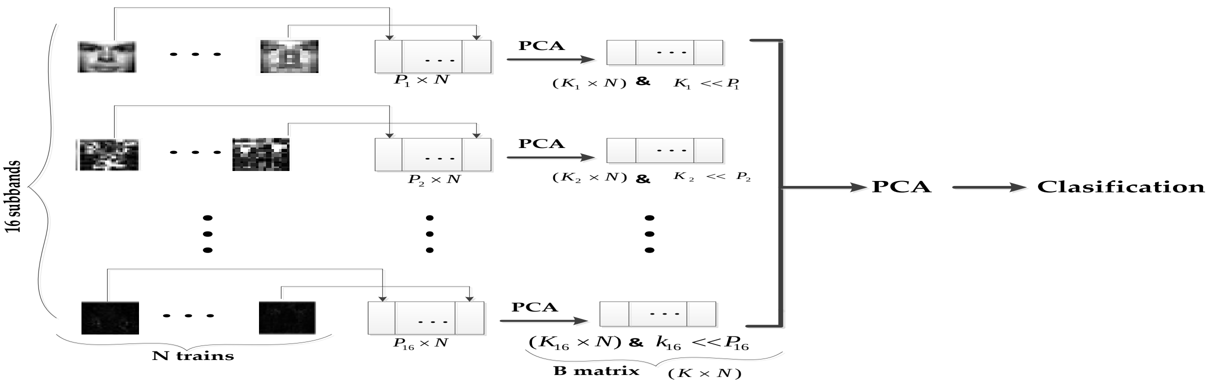

FIG. 5 THE BLOCK DIAGRAM OF OUR PROPOSED METHOD.

FIG. 6 SHOWN 28 SUBJECTS OF EXTENDED YALE B DATA BASE WITH SIZE 640×480 (TWO FIRST ROWS), AND THOSE MANUALLY CROPPED AND RESIZED TO 188×240 (TWO LAST ROWS).

Experimental Results

In this work, to implement the CT, the "pkva" filters (Do et al., 2002) were used and in order to evaluate the performance of our proposed method compared with some other methods, in addition, the extended Yale B face data base was used which contains 28 individuals under 64 different lighting conditions for nine poses. Since, in this paper the focus was on illumination variation, only frontal face images under 64 different lighting conditions were used. Therefore, $28\times64=1792$ images belonging to Yale B were used as test and train images. The original size of the images was 640×480 pixels. In our experiments, all images were manually cropped to include only the face with as little hair and background as possible and then resized to 188×240. Fig. 6 shows 28 subjects of extended Yale B face database and the corresponding cropped and resized images.

In addition, the facial images were divided into five subsets according to lighting angle ([$0^{\circ}\sim12^{\circ}$], [$13^{\circ}\sim25^{\circ}$], [$26^{\circ}\sim50^{\circ}$], [$51^{\circ}\sim77^{\circ}$] and greater than 77°) made with the camera's axis. So, five subsets were considered based on different lighting angles. These subsets for a sample image are shown in Fig. 7. In addition, the numbers of images to each subset are written in Table 1.

For these data base, half of the images per subset was selected for training and the remaining for test. Our experimental results are according to use 95% and 100% of subbands energy and the achieved accuracy are written in Table 2, where the average face recognition accuracy for 95% and 100% subband energy are 90.48% and 95.55% in order.

In other experiment, we considered only 6, 7, 8 images for each subject in each subset as training and the rest was used for testing, where 100% subbands energy was used and this experiment was performed on 5 subsets. The achieved recognition rate is written in Table 3.

The performance of our proposed method based on Error rate, in comparison with some other methods (Lee et al., 2005) is in Table 4, where, we used 8 images for each subject in each subset as training and the rest was used for testing; in which the error rate was obtained from the following equation,

$$\text{Error Rate} = \frac{|measured - actual|}{actual}\times100 \qquad (13)$$

Where measured is our result or experimental value and actual is accepted value or true value.

FIG. 7 FACIAL IMAGES OF A PERSON ARE CATEGORIZED INTO FIVE SUBSETS ACCORDING TO LIGHTING ANGLES.

Fig. 8. shows plots of error rates for four subsets of the extended Yale B database for comparison with (Lee et al., 2005).

On the other hand, the proposed method was compared to article (Hai-Long et al., 2010) in accordance with recognition rate. The results of this comparison showed the superiority of our algorithm. Recognition rate for different subsets of the extended Yale B database compared to other methods (Hai-Long et al., 2010) shows in Fig. 9.

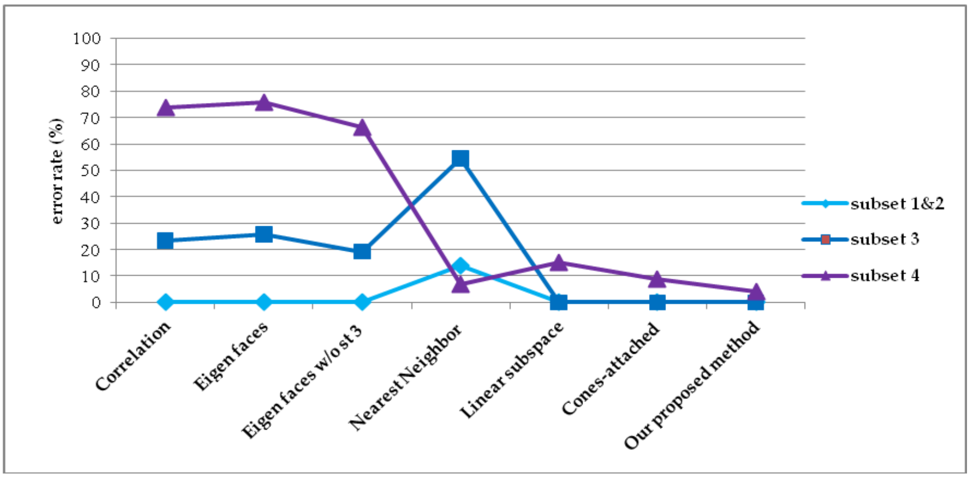

FIG 8 ERROR RATE IN DIFFERENT METHODS FOR FOUR SUBSETS OF THE EXTENDED YALE B DATA BASE.

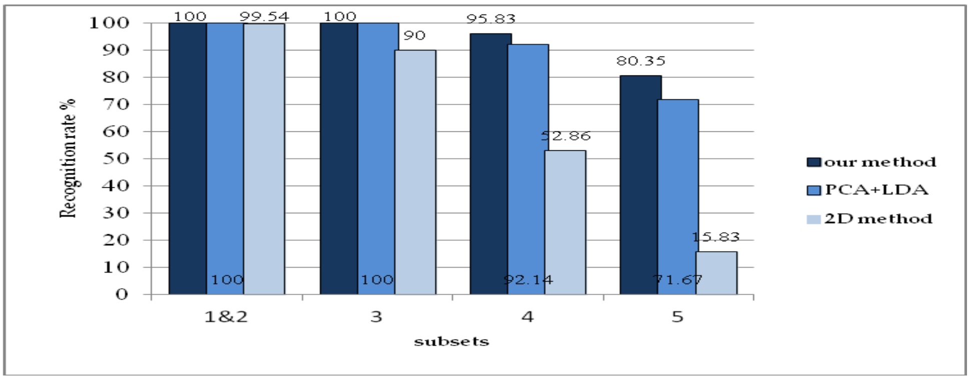

FIG 9 THE CHART SHOWS A COMPARISON BETWEEN OUR METHOD AND OTHER METHODS (HAI-LONG ET AL., 2010) FOR DIFFERENT SUBSETS BASED ON THE RECOGNITION RATE.

Conclusion

In this work, we have presented a face recognition system based on CT and PCA; and then decomposed the images into 3 scales and 15 orientations and used a total of 15 orientations along with the low pass subband. Further the PCA was applied twice on the subbands in order to reduce the dimension of input data and extract the feature vectors as well. In this case, all of the frequencies were used so our algorithm has relatively high accuracy. Experimental results on the extended Yale B database and in comparison with some other methods showed that our algorithm has appropriate performance.

TABLE 1 SUBSETS OF EXTENDED YALE B.

	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5
Lighting angle (deg)	[0 12]	[13 25]	[26 50]	[51 77]	>77
No. of images	$28 \times 7 = 196$	$28 \times 12 = 336$	$28 \times 12 = 336$	$28 \times 14 = 392$	$28 \times 19 = 532$

TABLE 2 RECOGNITION RATE OF OUR METHOD FOR EXTENDED YALE B DATA BASE.

	95% of subbands energy		**100% of subbands energy**	
	No. of CT coefficients	**Recognition rate (%)**	**No. of CT coefficients**	**Recognition rate (%)**
Subset 1	68	100	84	100
Subset 2	124	100	168	100
Subset 3	134	90,47	168	95,83
Subset 4	152	80,10	196	93,36
Subset 5	199	81,78	280	88,.57

TABLE 3 THE RELATIONSHIP BETWEEN THE NUMBER OF TRAINING IMAGES AND RECOGNITION RATE.

	Recognition rate (%)			
No. of training images	**Subset 1&2**	**Subset 3**	**Subset 4**	**Subset 5**
6	100	95,83	93,36	50
7	100	100	93,36	66,66
8	100	100	95,83	80,35

TABLE 4 COMPARISON OF DIFFERENT TECHNIQUES (LEE ET AL., 2005) BASED ON ERROR RATE ON THE EXTENDED YALE B DATA BASES.

	Achieved Error rate by different methods (%)						
	Correlation (Lee et al., 2005)	**Eigen faces (Lee et al., 2005)**	**Eigen faces w/o st 3 (Lee et al., 2005)**	**Nearest Neighbor (Lee et al., 2005)**	**Linear subspace (Lee et al., 2005)**	**Cones-attached (Lee et al., 2005)**	**Our proposed method**
Subset1&2	0,0	0,0	0,0	13,8	0,0	0,0	0,0
Subset3	23,3	25,8	19,2	54,6	0,0	0,0	0,0
Subset4	73,6	75,7	66,4	7,0	15,0	8,6	4,17

TABLE 5 RESULTS COMPARING THE PROPOSED ALGORITHM WITH OTHER METHODS (HAI-LONG ET AL., 2010) BASED ON RECOGNITION RATE ON THE EXTENDED YALE B DATA BASES.

	Recognition rate%				
Subsets	**1&2**	**3**	**4**	**5**	**Average**
2D method (Hai-Long et al., 2010)	99,54	90	52,86	15,83	71,56
PCA+LDA (Hai-Long et al., 2010)	100	100	92,14	71,67	92,76
Our method	100	100	95,83	80,35	95,23

REFERENCES

Adelman, Rachel. "'Such Stuff as Dreams Are Made On': God's Footstool in the Aramaic Targumim and Midrashic Tradition." Paper presented at the annual meeting for the Society of Biblical Literature, New Orleans, Louisiana, November 21–24, 2009.

Bamberger, R.H., and Smith, M.J.T. "A filter bank for the directional decomposition of images: theory and design," IEEE Trans. on Signal Processing. vol. 40, no. 4, 1992, pp. 882–893.

Burt, P.J., and Adelson, E.H. "The laplacian pyramid as a compact image code," IEEE Trans. Communications. vol. 9, no. 4, 1983, pp. 532-540.

Carstens, D.S., McCauley-bell, P.R., Malone, L.C., and Demara, R.F. "Evaluation of the human impact of password authentication practices on information security," Informing Science Journal, vol. 7,no. 1, 2004, pp. 67-85.

Do, M.N., and Vetterli, M. "Contourlets: a directional multiresolution image representation," In Processing. ICIP, vol. 1, 2002, pp. 357-360.

Georghiades, A.S., Belhumeur, P.N., and Kriegman, D.J. "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 23, no.6, 2001, pp. 643-660.

Hai-Long, Y., Huo-Rong, R., and Kai, Y. "Contourlet transform based algorithm of shadow compensation for face recognition," IEEE Intl. Conference. on Image Analysis and Signal Processing, 2010, pp. 671-674.

Kim, J.H., Park, J.G., and Lee, Ch. "Illumination Normalization for Face Recognition Using the Census Transform," SPIE Digital Library, vol. 6814, no. 681416, 2008, pp. 1-9.

Lee, K.Ch., Ho, J., and Kriegman, D. "Acquiring Linear Subspaces for Face Recognition under Variable Lighting," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 27, no. 5, 2005, pp. 684-698.

Mandal, T., Majmudar, A., Jonathan, Q.M. "Face recognition by Curvelet based feature extraction," International Conference on Intelligent Automation and Robotics, LNCS 4633, 2007, pp. 806-817.

Minh, N., Martin, V. "The Contourlet transform: an efficient directional multiresolution image representation," IEEE Transactions on Image Processing, vol. 14, 2005, pp. 1-16.

Moon, H., Phillips, P.J. "Computational and performance aspects of PCA-based face recognition algorithm," Perception. vol. 30, 2001, pp. 303-321.

Myers, B.A. "A brief history of human computer interaction technology," ACM interactions. vol. 5, no. 2, 1998, pp. 44-54.

Pirra, M., Gandino, E., Torri, A., Garibald, L., and Machorro-Lopez, J. M. "PCA algorithm for detection localization and evolution of damages in gearbox bearings," IOP Publishing, International Conference on Damage Assessment of Structures DAMAS, 2011, pp. 1-10.

Stonham, T.J. "Practical face recognition and verification with wizard," Aspects of Face Processing, Martinus Nijhoff Publishers, 1986, pp. 426-441.

Xie, X., Lai, J. and Zheng, W.sh. "Extraction of illumination invariant facial features from a single image using nonsubsampled contourlet transform," Journal of Elsevier, Pattern Recognition, vol. 43, 2010, pp. 4177-4189.

Yale B Yale university, http://cvc. yale. edu/ projects/. yale face B/ yale faces B. html, 2001.

Yang, L., Guo, B.L, NI, W. "Multimodality medical image fusion based on multiscale geometric analysis of Contourlet transform," journal Neurocomputing. vol. 72, no. 1-3, 2008, pp. 203-211.

Zabih, R., and Woodfill, J. "A non-parametric approach to visual correspondence," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996.